A COMPARATIVE ANALYSIS OF CNN AND TRANSFER LEARNING APPROACHES FOR MANGO LEAF DISEASE DETECTION

BY

SHAHRIAR EMON SHAYAM
ID: 201-15-3570

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Hasnur Jahan
Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Md Asaduzzaman
Senior Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2024

APPROVAL

This Project titled **"A Comparative Analysis of CNN and Transfer Learning Approaches for Mango Leaf Disease Detection"**, submitted by "Shahriar Emon Shayam", ID No: 201-15-3570 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on *26th January 2024*.

BOARD OF EXAMINERS

Dr. Sheak Rashed Haider Noori (SRH) — **Chairman**
Professor and Head
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Nazmun Nessa Moon (NNM) — **Internal Examiner**
Associate Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Dewan Mamun Raza (DMR) — **Internal Examiner**
Senior Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Dr. Md. Arshad Ali (DAA) — **External Examiner**
Professor
Department of Computer Science and Engineering
Hajee Mohammad Danesh Science and Technology University

DECLARATION

I hereby declare that, this project has been done by me under the supervision of **Hasnur Jahan, Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Hasnur 24-1-2024

Hasnur Jahan
Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

24-01-24

Md Asaduzzaman
Senior Lecturer
Department of CSE
Daffodil International University

Submitted by:

Shayam

Shahriar Emon Shayam
ID: 201-15-3570
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty God for His divine blessing makes me possible to complete the final year project/internship successfully.

I am really grateful and wish my profound my indebtedness to **Hasnur Jahan**, **Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of my supervisor in the field of "*Computer Science & Engineering*" to carry out this project. Her endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

I would like to express my heartiest gratitude to **Dr. Sheak Rashed Haider Noori**, Professor and Head**,** Department of CSE, for his kind help to finish my project and also to other faculty member and the staff of CSE department of Daffodil International University.

I would like to thank my entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, I must acknowledge with due respect the constant support and patients of my parents.

ABSTRACT

The comparative analysis of CNN and transfer learning approaches delves into the application of agriculture, with the specific focus on mango leaf disease detection. The core objective of the study is to develop an effective method for early, accurate and sustainable leaf disease detection system. With the help of convolutional neural networks and transfer learning techniques, this study aims to analyze raw mango leaf images to identify the certain disease. Raw images were collected from local mango orchards that undergo processing and data augmentation to refine model training. The study employs an array of deep learning models, including bespoke CNN models and pre-trained models such as EfficientNetB4 and InceptionV3. These models are conscientiously trained and tested against the dataset, with their performances evaluated on metrics like accuracy, precision, recall, and F1-score. It was observed that, the CNN-02 model, designed with batch normalization and dropout layers, exhibited superior performance in terms of accuracy and generalizability compared to other models, achieving 95.65% of accuracy. Additionally, the EfficientNetB4 model demonstrated an impressive learning capacity, with a precision rate of 98.3%. These results substantiate the effectiveness of both CNN and transfer learning approaches in the realm of leaf disease detection, with certain models showing exceptional accuracy and efficiency and promising potential of CNN and transfer learning techniques in early mango leaf disease detection, making a significant contribution to agricultural technology.

TABLE OF CONTENTS

CONTENTS | **PAGE**

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Agriculture plays an important role in the global economy. It contributes to 4% of the global gross domestic products (GDP), and also in some developing countries, it contributes more than 25% of GDP [1]. Being one of the largest industries in the world, agriculture employees more than a billion people and provides over 1.3 million dollars’ worth of food per year [2]. Among these crops, Mango (Mangifera indica), is a tropical crop popular in a lot of countries in the world. Not only it's very tasty, it is rich in vitamins, minerals and antioxidants which offers some health benefits including anticancer effects, improved immunity and digestive. [3]. It's cultivated in quite a few parts of the world, including Bangladesh being 9th country in producing mango [4]. Moreover, it has significant impact in some country's economy.

Due to the increasing spread of diseases in various crops, the global agriculture faces a lot of challenges. As a result, it causes significant economic losses and threats to food safety. One of the significant challenges is the detection and management of plant diseases. It can directly I have impact on crop quality. In case of identifying the diseases, traditional method of observing leaves manually, is often time consuming, labor intensive and most importantly, it can be inaccurate [5].

In this study, my primary goal is to use advanced techniques like convolutional neural network (CNN) and transfer learning to identify diseases in leaves by analyzing their images. These methods have shown promise in helping farmers detect diseases early and take the right steps. Throughout this process, this research explores and develops a method for early, accurate, and sustainable detection and diagnosis of leaf diseases. The following figure 1.1.1, 1.1.2, 1.1.3 are the examples of different types of leaf classes that I have aimed to classify using some transfer learning and convolutional neural network models.

Figure 1.1.1: Fresh Leaf

Figure 1.1.2: Powdery Mildew

Figure 1.1.3: Anthracnose

To achieve this, a collection of datasets with pictures of healthy mango leaves and leaves affected by common diseases are required. These images will be used to train the models about how to accurately identify different leaves with specific diseases. This study will also go through different symptoms and patterns of diseases in the leaves using data-driven approaches. To achieve the expected result, I plan to compare CNN and transfer learning approaches, which will examine the result for various pre-trained models to identify mango leaf diseases. The evaluation will consider factors like accuracy, computational cost, and how applicable these models are. This information will be beneficial for farmers and experts in accurately detecting and treating diseases. Farmers will be able to recognize diseases accurately and take actions in time to ensure the best quality of cultivation. This research doesn’t limit in mango leaves, it can be extended for any kind of leaf disease detection.

1.2 Motivation

Addressing the challenges faced by farmers and the agricultural industry has pushed researchers to speed up the study of diseases affecting mango leaves. Mango crops, which are crucial for many economies, are seriously threatened by diseases that harm food security and lead to significant financial losses.

Traditionally, leaf diseases are checked just by inspecting the leaves visually. The process is slow and not accurate enough as the process is physical. Using the traditional methods can result in delayed or ineffective treatment, especially when the diseases are needed to be detected as soon as possible. That's why more advanced, automated, and environmentally friendly techniques are needed that can spot diseases on leaves quickly and accurately.

Farmers are focusing on more sustainable and eco-friendly ways of farming. One important aspect is managing diseases properly. It's important because it helps protecting the environment as well as reducing the need for harmful chemicals and ensures that mango farming can thrive in the long run. With the help of modern technologies like computer vision, machine learning, and image processing, we have an opportunity to make a big difference in how we detect diseases on mango farms. These technologies allow us to look at a lot of data, find patterns related to diseases, and teach computers to accurately spot and diagnose diseases that affect mango leaves. This can be a game-changer for mango farmers and the environment.

The research aims on certain domains like healthy farming, protecting crops, and making the farming industry stronger. It's not just about improving mango farming, it's also about finding better ways to deal with diseases in other crops as well. The aim is to provide farmers with a reliable tool to combat diseases, increase crop production, and promote environmentally friendly farming practices.

1.3 Rationale of the Study

The rationale of this study arises from the need to find better ways to deal with problems in the agriculture field, especially when it comes to growing mangoes in Bangladesh.

Disease Impact on Crop Yield:

Mango plants can get affected by various things like virus, fungus. bacteria or other pests, and it greatly affects the quality of fruit and quantity of production. The usual methods of inspecting the leaves visually requires a lot of time, also it can be inaccurate in practice. This research aims to use smart computer techniques to quickly and accurately spot diseases, which will help identifying diseases faster and avoid losing crops.

Advancements in Deep Learning:

In this research, I aim to use a modern-day computer technology called convolutional neural network to identify leaf diseases by processing the leaf images. These techniques are developed for recognizing objects. I will also use transfer learning in this research, which is also included in CNN. But instead of creating the models from scratch, these models are already trained on certain dataset and can be applied on other training process. The goal is to use these smart techniques to make models that can find and name mango leaf diseases in time accurately.

Resource Optimization:

In regular farming, farmer usually check the crops by hand which takes a lot of time and effort, especially when there's a lot to grow. By implementing this research to its potential, it can make things easier by using automated models that can spot diseases in the crops. This research aims to use CNN and transfer learning for identifying the crop diseases which will guarantee a more accurate and thorough evaluation of the crop's health while making the best utilization of labor and time assets.

Comparative Analysis for Model Selection:

Using the deep learning methods directly into application is not enough. There are numerous techniques and models, each of them having individual processing technique, which generates different results for different situations. The comparative analysis between CNN and transfer learning model will help us figure out which one is best for detecting diseases in mango leaves. It's important to know what is the strength and weakness of every model so we can use the right one for future applications in agriculture.

Technology Adoption in Agriculture:

In this digital era, we can see the application of modern technologies in agriculture sector in many developing countries including Bangladesh. This study also contributes in encouraging the use of technology in farming, which helps the vision more modern and sustainable. The results of the research can help policymakers and stakeholders make smart choices about bringing new technologies into regular farming.

Social and Economic Implications:

Beyond the technical aspects, the study recognizes the broader socio-economic implications of improved disease detection. By safeguarding crop health, the project aims to contribute to food security, support farmer livelihoods, and reduce economic vulnerabilities in rural communities. The outcomes of this research have the potential to foster a positive impact on both individual farmers and the agricultural community at large.

In summary, the rationale for this study lies in the intersection of agricultural challenges, technological advancements, and the potential for positive socio-economic impact. By addressing the specific needs of mango cultivation in Bangladesh through a comparative analysis of CNN and transfer learning, this research seeks to pave the way for more resilient, efficient, and technology-driven agricultural practices.

1.4 Research Questions

There are frequently asked questions that might arise regarding this research. Such question includes:

1. Is there any research done on mango leaf diseases in Bangladesh?
2. How did they figure out what illnesses were affecting the mango leaves?
3. Is there any dataset available regarding the topic? How big is the dataset?
4. What kinds of diseases and issues are grouped into different categories?
5. In what ways do these methods and techniques benefit for the society?

1.5 Expected Outcomes

The main objective of this study is to determine the most effective method to identify diseases affecting mango leaves. This study revolves around comparing the effectiveness of convolutional neural networks (CNN) and transfer learning approaches in achieving accurate and reliable detection of mango leaf diseases. Expected results include careful evaluation of model performance, taking into account metrics such as accuracy, efficiency, and adaptability to different conditions. By analyzing the strengths and limitations of every method, this study provides practical insights to help researchers and experts choose the most appropriate approach for automatic detection of mango leaf diseases. The aim of this study is not only to provide new knowledge in the field of image processing in agriculture, but also to validate these results by comparing them with existing research results to improve and further develop mango leaf disease detection methods that will pave an opportunity for future research.

1.6 Project Management and Finance

To implement the recommended models, a decent amount of raw images were required. I have tried to collect as much data as possible for better result. All of the images were collected using my smartphone. The collected data underwent training with several pre-trained Keras models, with the help of Google Colab, which provides extra computational power without requiring special hardware. This choice enables uninterrupted and efficient model training. It's important to highlight that, during the training phase, the models in our study underwent different numbers of training cycles, known as epochs or iterations. This variation was crucial to assess how the models performed over time and to find the optimal balance between learning from the data and preventing overfitting. This approach ensures that our models are well-prepared to detect mango leaf diseases accurately. Moreover, I have managed the project really well, organizing everything carefully. Plus, collecting data didn't cost anything.

1.7 Report Layout

In this report, I have been through several chapters of discussions. Starting with Chapter 1, where I have tried to establish the importance of what I am doing. I have highlighted the significance of mango regarding this research, my motivation that inspired me to do this research, the final outcome that I expect from this study and a brief overview of my report. Moving on to the chapter 2, where I have studied the related works of what I am doing. It also helped me to identify what are the flaws of the previous study, and the loopholes that I can improve in my research. This chapter also discusses the challenges that I might face throughout the research. chapter 3 of this study includes the process of my research. Starting from data collection to model implementation, I have illustrated all the preparations of my research in this chapter. After that, I have discussed the computational processes and the results of my research in chapter 4. It describes the models, how those models worked in my dataset and what actually happened in the training process. Finally, I added the conclusion part of my research in chapter 5 where I have discussed the impact of my study in society as well as environment, also the future potential my research holds.

CHAPTER 2
BACKGROUND

2.1 Preliminaries/Terminologies

Researchers have been through different processes to find a conventional way of recognizing leaf diseases. Machine learning and deep learning offers such methods that can identify patterns from the symptoms of leaves and understand and predict those diseases. There are numerous models that are used to identify leaf diseases, with each model performing differently according to the trained dataset. There are numerous researches involves with the leaf disease prediction, each having individual advantage, as well as limitations. In this chapter, other research works will me summarize in short, which will provide a proper navigation to guide this research work.

2.2 Related Works

A lot of work has been made in using deep learning to find plant diseases, and many researchers have made important contributions. On our way back, we see how far this field has come and how much it has changed.

The story starts with the groundbreaking work of Mohanty et al. (2016) [6], who used deep convolutional neural networks and transfer learning in new ways. Their groundbreaking work on the gray scaled PlantVillage dataset got it right 98.21% of the time, which is very good. But the model's lower accuracy in different situations showed that it needed to be made more stable and flexible.

In 2017, Bin Liu et al. [7] and Yang Lu et al. [8] added new parts to the story. Bin Liu et al. studied apple leaf disease and got a 97.62% success rate while reducing model parameters by a lot. However, their work was hindered by the lack of diseased images they had access to. At the same time, Yang Lu et al. looked into how to identify diseases in rice using digital image processing techniques in a CNN-based design. Their creative method worked 95.48% of the time.

After that, finding diseases on mango leaves became the main topic. Singh et al. (2017) [9] studied anthracnose disease. With MCNN, data preprocessing and picture enhancement, they were able to get a 97.13% success rate. The years 2018 saw more progress in this area thanks to the work of Merchant et al. (2018) [11] and

Arivazhagan and Ligi (2018) [10]. Arivazhagan and Ligi suggested a CNN method that worked 96.67% of the time, and Merchant et al. looked into feature extraction through digital picture processing.

As I looked into it more, I found interesting work that wasn't just finding plant diseases in fruits and veggies. Chen et al. (2019) [12] did a great study that presented "LeafNet," a new CNN architecture made to find diseases in tea leaves. With a 90.16% accuracy rate, LeafNet did much better than standard algorithms like SVM and MLP. In the same way, Agarwal et al. (2019) [13] made a lot of progress with their CNN model for finding diseases on tomato leaves. By adding more data to improve class balance, they were able to get their model to a respectable 91.2% accuracy across a wide range of disease types, beating even pre-trained models like VGG16, InceptionV3, and MobileNet.

Another important study from this year is Aravind et al.'s (2019) work on the disease that affects Solanum melongena (eggplant) [14]. Images of plants, either whole or partially, were used to make the collection. Disease classification used deep learning models that had already been taught. VGG16 models were more accurate than AlexNet models. The better VGG16 model got an average validation accuracy of 96.7%, and on field sample shots, it got an accuracy of 93.33%.

On the other hand, Arya and Singh (2019) used both potato and mango leaves in a comparative analysis study on neural networks and transfer learning for disease identification [15]. They changed the size of her dataset and then added to it using affine transformation, perspective transformation, picture rotation, and intensity transformation, among other methods. After that, this bigger set of data was used to teach both the CNN model and the Alexnet model. The results showed that AlexNet did much better than the CNN model when using a pre-trained transfer method. It was able to correctly identify 99.75% of diseases.

The 2019 study by Trang et al. looked into mango leaf disease using the AlexNet model along with other well-known deep learning models like Inceptionv3 and MobileNetv2 [16]. The result shows that the Inceptionv3 models, more specifically the other two, with accuracy of 78.48%.

In the same vein, Venkatesh et al. [17] and Maheshwari and Shrivastava [18] (2020) added to the field of mango leaf disease classification by using transfer learning and data enrichment to make their work more accurate.

Pankaj Kumar et al. [19] came up with a new CNN design for finding anthracnose disease in mango leaves in 2021. It worked 96.16% of the time. In the same way, Rajbongshi et al. [20] looked at different CNN designs and found that InceptionV3 was the most accurate, at 96.67%.

Sharma et al. [21] created a new deep-learning system in 2022 that could tell if mango leaves were sick. As part of their method, they skillfully improved training pictures by rotating, scaling, and even separating them into groups based on the patterns of leaf veins. They used the power of a Convolutional Neural Network (CNN), which is known for being good at recognising images. With an amazing 90.36% accuracy, this complex model was able to identify different diseases on mango leaves, showing a big step forward in using advanced AI to find plant diseases.

The newest developments in 2023 are Jyothi and Kranthi's use of DenseNet201 [22] and Rizvee et al.'s introduction of LeafNet [23]. Both of these achieve ground-breaking accuracy in finding diseases on mango leaves (with an average accuracy of 98%). Mimi et al. [24] made an Android app for finding diseases on rubber leaves by adding pictures of Catharanthus roseus and strawberry leaves to a dataset. They got the best accuracy of 97.35% by using a mix of hybrid CNN and SVM, CNN, and MobileNetV2-based transfer learning models. Vijay and Pushpalatha [25] trained different models, such as CNN and EfficientNetB4, and got an impressive 93.01% accuracy.

2.3 Comparative Analysis and Summary

The comparison of study visualizes that, AlexNet, MobileNetV2, VGG-16 are the commonly used transfer learning models for most of the studies, achieving successful results in their applications. That shows the possible scope of application for other transfer learning-based model. The following table 2.3.1 shows the summary of some related works of this research.

TABLE 2.3.1: SUMMARY OF RELATED WORKS

First Author Name	Area of Studies	Models	Accuracy
Vijay C. P.	Mango Leaf Disease	Hybrid CNN with EfficientNetB4, VGG-16	93.01%
Krishnaswamy R. Aravind	Solanum Melongena Disease	Modified VGG-16, AlexNet	96.7%
Aditya Rajbongshi	Mango Leaf Disease	CNN, DesNet201, InceptionResNetV2, InceptionV3	96.67%
Afsana Mimi	Rubber Leaf Disease	Vanilla CNN, CNN-SVM Hybrid Model, MobileNetV2	97.35%
Uday Pratap Singh	Mango Anthracnose Disease	Multilayer CNN	97.13%
Sunayana Arya	Potato and Mango Leaf Disease	CNN and AlexNet	98.33%
Pankaj Kumar	Mango Anthracnose Disease	Deep CNN	96.16%
Jing Chen	Tea Leaf Disease	LeafNet	90.16%
S. Arivazhagan	Mango Leaf Disease	Custom CNN	96.67%
Kien Trang	Mango Leaf Disease	InceptionV3, AlexNet, MobileNetV2	88.46%

2.4 Scope of the Problem

This research aims on several domains in agricultural technology. Initially, this research inspires to pinpoint agricultural challenges, with a specific focus on promptly alerting field-owning farmers about the presence of early or late rot through my implementations. Achieving this goal required a systematic approach, though not without its challenges.

One notable domain is the overlapping information between powdery mildew and sooty mould. The two categories are so identical that it becomes challenging to differentiate between them accurately. Additionally, some foliage may carry a mild infection that is not easily visible, potentially leading to a misdiagnosis, as my approach might be mistaken for various other illnesses.

Through the use of a Convolutional Neural Network (CNN) and the EfficientNetB4 model predictor, I have demonstrated the ability to analyze images of mango leaves with an impressive accuracy result. This example provides us with optimism about the potential success of the methodology. In summary, this research delves into agricultural issues, aiming to provide farmers with timely alerts about rot on their fields. While the methodical approach is promising, challenges arise from the overlapping characteristics of different diseases. The successful application of CNN models to analyze plant images further encourages the potential effectiveness of this approach.

2.5 Challenges

Every research implementation involves a lot of challenges. Using deep learning algorithms for disease classification, models are trained to understand patterns and information from diseased leaves. Each disease involves unique symptoms and visuals. Some diseases can be identified by spot. Some of them can be identified due to lack of nutrients. Also, pests, fungus and viruses are involved in many diseases. Sometimes, there can be multiple disease symptoms on a single leaf. It is difficult to identify the disease in some cases, without a proper supervision of any expert. In another aspect, diseases and pests are not the only cause of symptoms. Sometimes the surrounding circumstances and local environment can manifest different symptoms. As the image processing is the key of deep learning method, it is hard to correctly identify disease just by visual inspection.

Another challenge in identifying the leaf disease lies with the functionality of image processing. There are different ways to teach the computer to recognize these diseases. Some methods work better in certain situations and dataset, but it's tough to know which one is the best for mango leaves. Plus, using a computer to spot diseases needs a lot of data and powerful technology. Also getting the right kind of data can be tricky because mango fields are spread out, and diseases vary in the change of seasons and how common they are in different places.

Also, comparing different computer methods means spending a lot of time and effort. It's more like trying to figure out which tool works best for the job. But theoretically, it's hard to get all the algorithms that is needed for the specific research because they might be computationally complex and difficult to implement. And when you're working with plants, things like bad weather or changes in the seasons can make the job even tougher.

So, even though finding diseases on mango leaves using computers sounds cool, it comes with its share of challenges, like figuring out the best way to teach the computer, getting the right data, and dealing with different tools and environmental obstacles.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Research Subject and Instrumentation

My research mainly focuses on figuring out different diseases that affect mango leaves. In this research, some neural networks from deep learning methods will be applied on a dataset, which will help the model to train different disease patterns so that the model can be used for efficient disease prediction. To conduct this research, a lot of pictures is needed that will help the model understand different disease symptoms. In my study, I have used image processing techniques to group the diseases into three main categories: Anthracnose, Powdery mildew, and healthy leaves. I went to two orchards to take real-time pictures using my Google Pixel 2XL for data collection.

To handle the computer part of my research, I used the Python programming language which involves other tools like NumPy, Scikit-Learn (SkLearn), and OpenCV. All the testing and training stuff was done on a Windows computer, and I used Google Colab to do the analysis. It was a systematic process to make sure everything works smoothly.

3.2 Data Collection Procedure

For any research project, it is critical to train the models correctly. The training process required a lot of data. A model's learning ability and efficiency is depended on the scale of the dataset. Highest number of images will help the model to achieve the best training and prediction performance. I obtained the information I needed from local field. The data collection process led me into two local mango orchards in my neighborhood. I took around 300 photographs using my Google Pixel 2XL smartphone in Rajshahi's diverse fields, specifically from the orchards. The data collection from multiple locations was difficult, time consuming for a single person but it was completed. The following figure 3.2.1 shows a brief visual of the dataset.

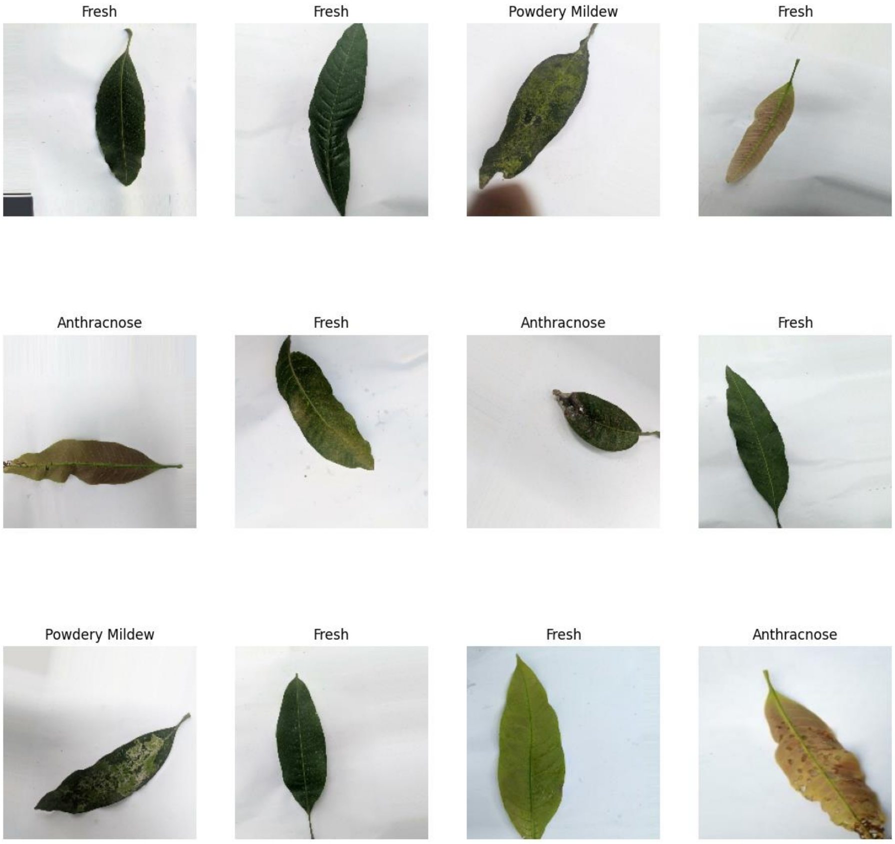

Figure 3.2.1: Dataset

3.3 Statistical Analysis

Using mango leaves, I generated distinctive sets of images representing three different classes in my study. These conditions were associated with Anthracnose, powdery mildew, and healthy leaves. I divided the dataset into training, testing, and validation sets to evaluate the model's effectiveness. In the training set, every image was used, while 15% were used for testing and another 15% for validation. I can offer more specific information because I have a collection of 1741 images at my disposal. The dataset was expanded to improve accuracy, which contributed to the increased dataset size. The following figure 3.3.1 illustrates the dataset ratio of the classes.

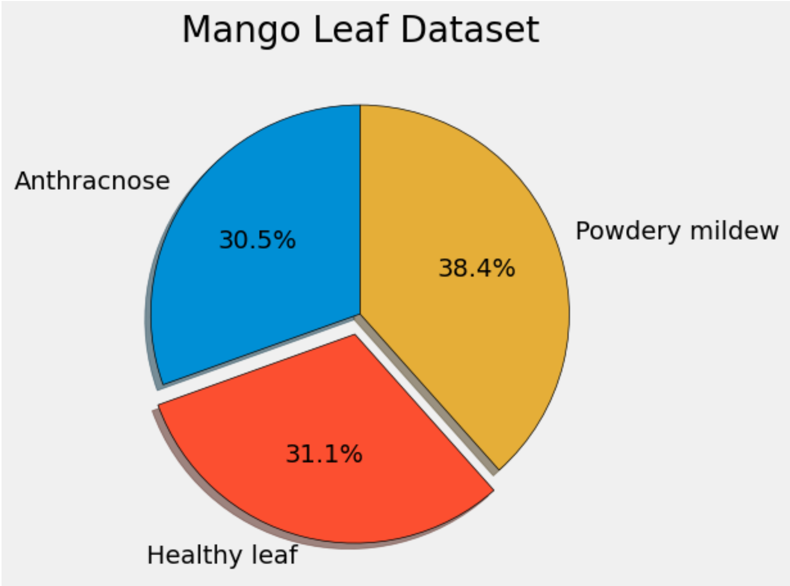

Figure 3.3.1: Dataset Statistics

3.4 Proposed Methodology

A proper planning and road mapping is the key to every successful research implementation. The base of this research involves the study and implementation of deep learning. Deep learning is a super powerful way to use computers. It can handle a lot of information, especially in image processing. CNN stands for "Convulsive Neural Network," and it's commonly used for processing images. Back in the 1980s, people started studying CNN. This special kind of network is based on how the neurons in our visual cortex, or the part of our brain that sees things, work. It's different from the regular networks we use. The name "convolution" comes from math and means combining two things to make a third thing. It's like a mix of two inputs, and it can also be called enhancement.

Transfer learning, a branch of machine learning in which a model created for one task is later used for another. The idea is to get the model used to predicting something for a particular task. Using a model that has already been applied to a new task is the fundamental principle of transfer learning. In fact, in this transfer learning, we often attempt to transfer information in as many different ways as possible from the prior task the model was trained on to the current task at hand. Data can be changed using an approach called image augmentation. The following figure 3.4.1 shows the key roadmap of the research process, which is known as methodology.

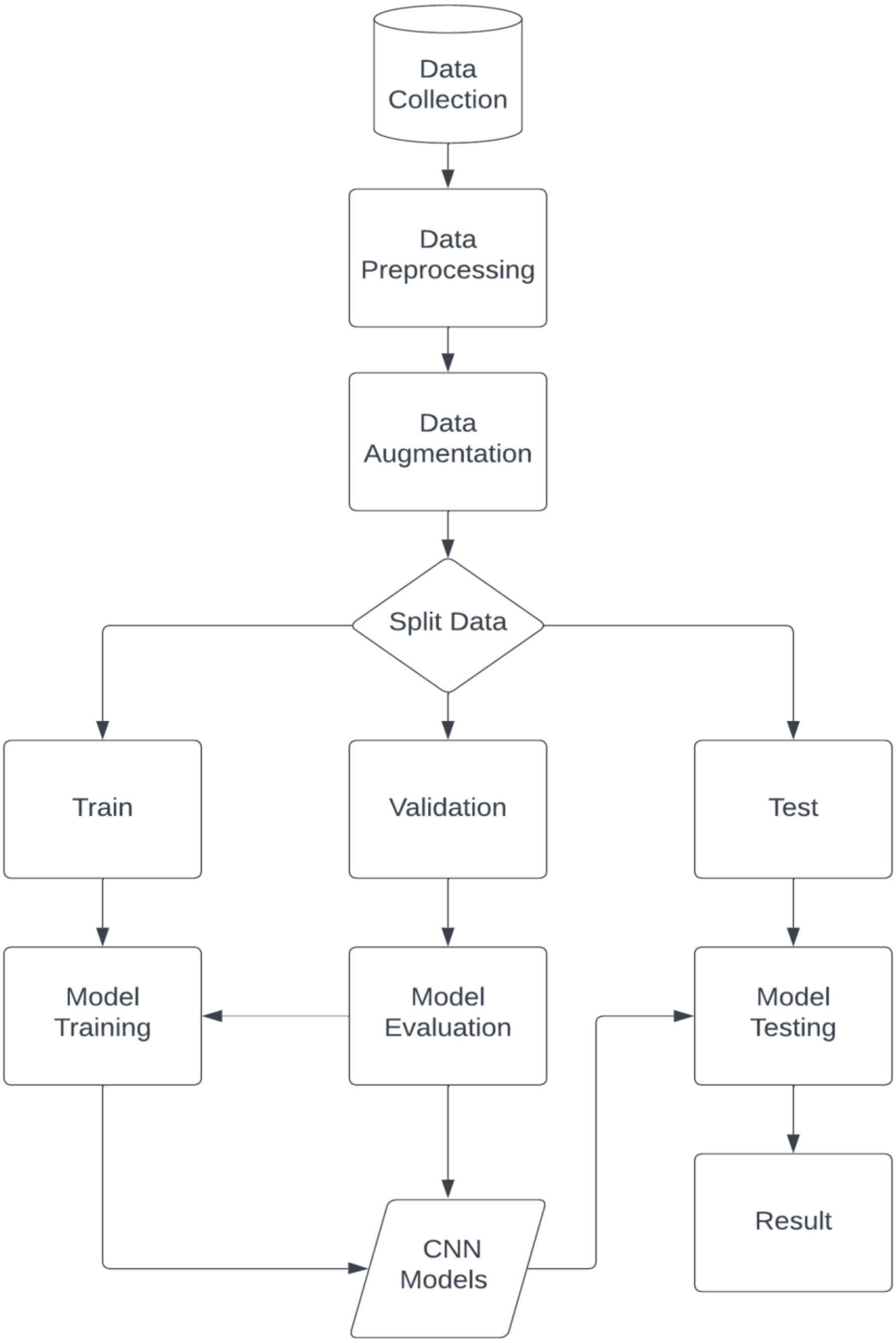

Figure 3.4.1: Methodology

To conduct this research, first, the data collection process will be completed. After collecting images by properly identifying the diseases, the dataset will be preprocessed for better model training. As the size of dataset, number of images and other definitions can hamper the model training process, it is necessary to process the dataset. After that, variety of deep learning models will be employed to find mango leaf illnesses in the dataset. The phases of the approach are data collection, preprocessing, data augmentation, training model including CNN, inception V3, EfficientNetB4, and model testing. Once the model's pre-training was completed, I assessed its accuracy by comparing it to the test dataset.

Data Augmentation

I collected some raw images from different sources for the research. But raw images cannot be deployed for model training without properly adjusting the image definitions. So, image augmentation is the solution for insufficient data. Image augmentation increases the number of images, which helps the model to train better and produce better result. Using image rescale, the pictures were reduced in size to have a resolution of 240x240. I have also applied rotation, width and height shifting, shearing, zooming, and horizontal flipping for augmentation. The dataset now contains roughly 1741 photos after augmentation, making it appropriate for deep-learning model implementation.

Applied Models

In deep learning, Conv2d is used in the models to quickly calculate the learnable layer, also known as the convolutional layer in a transfer learning model, which effectively improves the classification size of the input images. It essentially works by combining the results of depth wise convolution on pointwise convolution. On the other hand, the primary purpose of MaxPooling2D is to reduce the size of the required input photos. The model is then Flattened into a one-dimensional representation, the photos are fitted with Dense, and the model is primarily classified with Dense.

After successfully applying preprocessing and data augmentation on the dataset, I have gone through some custom and pre-trained models. Convolutional Neural Network offers a huge number of models depending on how they perform in certain datasets. In my research, I have tried to build custom models along with using pre-trained models.

CNN-01

In my study, the model I chose includes Convolutional Neural Networks (CNNs). CNNs are used for ability to automatically learn and extract hierarchical features from input data, making them well-suited for tasks like image recognition and classification. To construct the CNN model for my research, I started with a Convolution Layer using 32 different filters and the best Kernel Size of [3x3]. I added a special activation function called rectified linear unit (ReLU) to this layer. After that, I put a max pooling layer with a [2x2] size to make things simpler. I kept building more layers using the same convolutional and pooling idea, increasing the depth and filter size to 64, 128, and 256.

To make everything work together, I added a flatten layer after all the convolutional layers. This layer turns the output of the convolutional layers into a flat vector. Then, I made the dense neural network layers, starting with one layer having 512 units and using ReLU activation. Finally, for the last result, I added another layer with a softmax activation function for dealing with different categories. This last step is super important to get the right classification for training the model successfully.

CNN-02

Another CNN model was used in this research that is mostly inspired from the previous CNN model, but revised with the addition of Batch Normalization and dropout layers. The model has the primary convolutional layer of 32 filters and ReLU activation, followed by a max pooling layer with a 2x2 pool size and a batch normalization layer. Batch normalization enhances the training stability of the model and convergence of neural networks by normalizing input layers. Using the similar convolutional and pooling layer, multiple layers with increasing depth and filter size of 64, 128 and 256 were constructed subsequently with batch normalization. For the construction of the dense layers, dropout layer is added after the primary dense layer, which drops half of random neurons during each training, which prevents the model from overfitting and enhances the model with robust learning of features by avoiding reliance on specific neurons. The architecture of the model is described in the following figure 3.4.1.

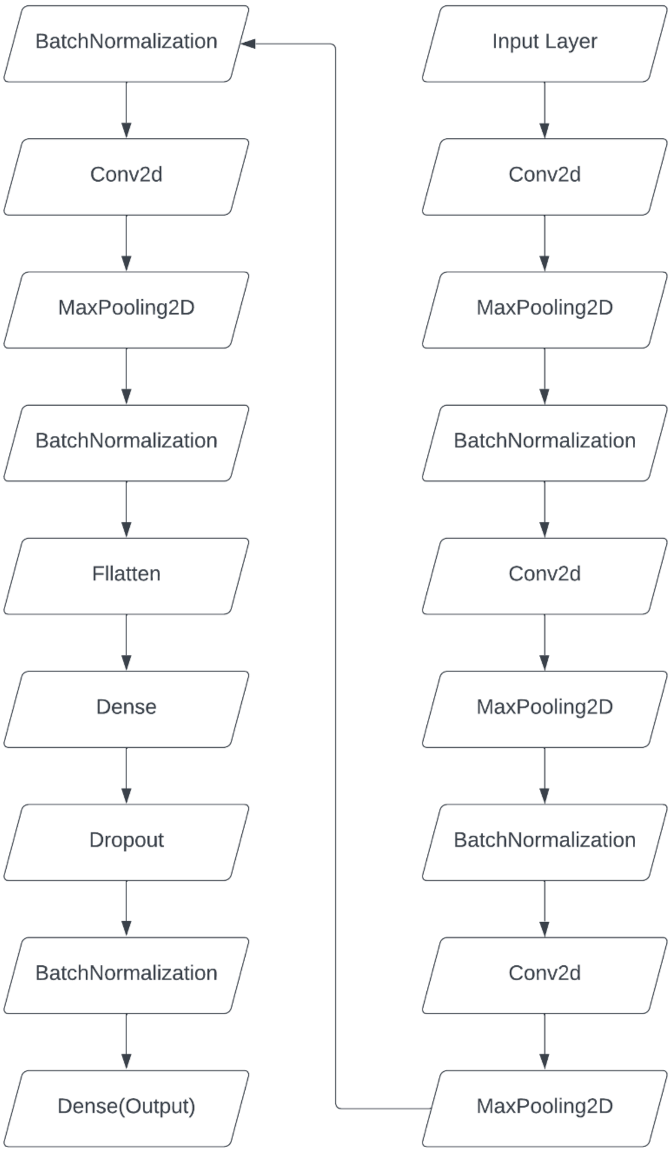

Figure 3.4.1: CNN-02 Architecture.

EfficientNetB4

EfficientNetB4 is a smart kind of computer model designed to be just the right size and work really well. It comes from the earlier EfficientNet family. This model blocks of different techniques like batch normalization, convolutional layers, and activation methods such as Swish or ReLU. The trick with EfficientNetB4 is to keep things balanced. It does this by making the network wider and deeper at the same time. The depth scaling happens by adding more layers to the network structure. To make it wider, the number of channels in each layer is adjusted, thanks to the compound scaling, it not only makes the network deeper or wider, but also it considers both things together.

EfficientNet offers immense performance while utilizing fewer resources than classic designs by prioritizing parameters that have a major impact on model correctness. The model used in the research offers freezing of the layers of the model so that their weights don't get updated during model training. A layer of a global average pooling, followed by the input layer, is used to reduce the spatial dimensions of the input to a single value for each feature map. The final touch to the model is applied with a dense output layer with softmax activation for the final classification. The overall model summary is shown in the following figure 3.4.2.

```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_layer (InputLayer)    [(None, 240, 240, 3)]     0

 sequential_1 (Sequential)   (None, 240, 240, 3)       0

 efficientnetb4 (Functional  (None, 8, 8, 1792)        17673823
 )

 global_avg_pool_layer (Glo  (None, 1792)              0
 balAveragePooling2D)

 output_layer (Dense)        (None, 3)                 5379

=================================================================
Total params: 17679202 (67.44 MB)
Trainable params: 5379 (21.01 KB)
Non-trainable params: 17673823 (67.42 MB)
_________________________________________________________________
```

Figure 3.4.2: Model Summary of EfficientNetB4

Inception V3

InceptionV3 is a convolutional neural network architecture designed for image classification and object detection tasks. It is known for its inception modules, which capture features at different scales within the same layer, resulting in efficient and accurate neural network training from images. Inception model in transfer learning, serves as a feature extractor after being trained on a big dataset for general image recognition. InceptionV3's first layers, which have learned general features, are frozen, and further layers are introduced to adapt the model to a particular task.

In this research, the InceptionV3 model is fine-tuned by freezing the weights of initial layers and adjusting the final layer. The last 15 models are allowed to be trainable while the initial layers remain frozen. A global average pooling layer, followed by a flattening layer, and two densely connected layers with 512 units each and ReLU activation are applied in the model. Between these dense layers, dropout layers are added for regularization. The final dense layer has a softmax activation function for multiclass classification. The overall model summary is shown in the following figure 3.4.3.

```
 global_average_pooling2d (  (None, 2048)                0         ['mixed10[0][0]']
 GlobalAveragePooling2D)

 flatten_2 (Flatten)         (None, 2048)                0         ['global_average_pooling2d[0][
                                                                   0]']

 dense_4 (Dense)             (None, 512)                 1049088   ['flatten_2[0][0]']

 dropout_1 (Dropout)         (None, 512)                 0         ['dense_4[0][0]']

 dense_5 (Dense)             (None, 512)                 262656    ['dropout_1[0][0]']

 dropout_2 (Dropout)         (None, 512)                 0         ['dense_5[0][0]']

 dense_6 (Dense)             (None, 3)                   1539      ['dropout_2[0][0]']

==================================================================================================
Total params: 23116067 (88.18 MB)
Trainable params: 1708163 (6.52 MB)
Non-trainable params: 21407904 (81.66 MB)
__________________________________________________________________________________________________
```

Figure 3.4.3: Model Summary of InceptionV3.

3.5 Implementation Requirements

For the computational process of this research, the system I have used includes

- Personal Computer
- Operating System: Windows 10
- File Management: Google Drive
- Programming Environment: Google Colab
- Programming Language: Python with Tensor Flow

CHAPTER 4
EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Experimental Setup

In this experiment, the raw images were collected manually using smartphone, then the images were uploaded into the google drive. For the programming environment, Google Colab was selected as it offers intense computational power of high-end processor and graphics processing unit for free, using the local browser only. By simply importing and installing the required tensor packages and keras pre-trained models for our experiment, it is ready to process the data. Using the python programming language, the preprocessing and data augmentation were done and then the data were used for the model training.

4.2 Experimental Results and Analysis

Using my dataset, I assessed the precision, recall, and F1 score to determine the efficacy of the design in this study. I use acronyms like True Positive and False Positive to express the results. Similarly, the terms True Negative and False Negative are interchangeable. A model's performance is determined by analyzing the training and validation accuracy, training and validation loss graph, and the accuracy of the model using the confusion matrix.

Analyzing a model result through training and validation metrics includes the plotting and observing the graph over the training epochs or iterations of model compilation. The graph helps with convergence in training accuracy and loss, helps identifying whether the model is overfitting or underfitting. In practice, the model compilation can be stopped by analyzing bad training loss and accuracy, which can be helpful for efficient model compilation and save time. Training accuracy and loss highly depends on the number of epochs or iteration.

Confusion matrix is also important asset that can evaluate a model's performance, especially for classification models, providing a detailed breakdown of the training and validation process of a model. Confusion matrix is also crucial for imbalanced dataset training, which will be the optimal evaluation asset for this research. Confusion matrix has some basic components that helps with analyzing a model's performance.

True Positives (TP): Instances in which the model predicts the positive class properly.

True Negatives (TN): Instances in which the model predicts the negative class properly.

False Positives (FP): When the model predicts a positive class but the true class is a negative class (Type I error).

False Negatives (FN): Situations in which the model predicts a negative class while the true class is positive (Type II error).

Based on these instances, confusion matrix generates results of a model in different scales. These are precision, recall, f1-score, accuracy.

Precision: Precision is the ratio of true positive predictions to the total number of instances predicted as positive including both true positives and false positives. It measures the accuracy of positive predictions. A high precision states that the model predicts positive instances accurately. Precision is described as

$$Precision = \frac{TP}{(TP + FP)}$$

Recall: Recall, also known as sensitivity is the ratio of true positive predictions to the total number of actual positive instances including both true positives and false negatives. It measures the ability of the model to capture all positive instances. A high recall states that the model is effective at identifying most of the positive instances. The formula of recall is

$$Recall = \frac{TP}{(TP + FN)}$$

F1-Score: F1 score is the harmonic mean of precision and recall. It provides a balanced measure of precision and recall which is useful when there are imbalance classes in dataset. F1 score ranges between 0 to 1, where a higher value indicates a better balance between precision and recall. F1-score is calculated as

$$F1\,Score = \frac{2 * (Precision * Recall)}{(Precision + Recall)}$$

Accuracy: Accuracy is the ratio of accurate predictions of true positives and true negatives to the total number of instances. Accuracy provides an overall measure of how well the model is performing across all classes. However, it may not be suitable for imbalanced datasets. Following equation determines the calculation of accuracy:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

To check how well my model is doing, I have used terms like "true positive," "true negative," "false positive," and "false negative." These help me figure out how accurate my model is performing. I find them really useful to evaluate the results. I'm going to share graphs of training and validation accuracy and loss of all models. These graphs illustrate how the lines go up and down. If the lines fluctuate a lot from each other and the end points of validation accuracy and loss don't match, it means the model might not be good. But if the lines are steadier and the end points are close to each other, then it indicates that my model performed well. This way, it's easy to understand how my model is performing.

Accuracy & Loss Graph

There are several key factors that helps us evaluate the performance of a deep learning models which includes validation accuracy and validation loss. Validation accuracy visualizes how often a model makes correct predictions on new, unseen data. On the other hand, validation loss illustrates the degree of incorrect predictions made by the model in predicting validation images. High validation accuracy and low validation loss indicate a model's better performance at learning and making accurate predictions on unfamiliar data. So, to evaluate the performance of the models used in this research, we have to observe these metrics during model training to decide when to stop training. Otherwise, the models will face issues like overfitting or underfitting.

Figure 4.2.1: Accuracy and Loss Graph of CNN-01

Here is the graph of the CNN-01 model shown in figure 4.2.1 that is used in our research which describes some interesting information. If we inspect the graph closely, initially, the training accuracy is high and the validation accuracy is low, indicating overfitting. This is because the model is able to perfectly learn the training data but performs poorly on unseen validation data. But as the training progresses, both training accuracy and validation accuracy increase, indicating better performance. Eventually, both curves level off and become very similar. This is a sign of good generalization and performance on both the training and validation datasets.

The graph of the training and validation loss shows a similar trend. Initially, the training loss is low and the validation loss is high, indicating underfitting. As the training progresses, both training loss and validation loss decrease, indicating the model is learning from the data and improving its performance. In summary, the provided graph indicates a well-performing CNN model that generalizes well to both the training and validation datasets. This suggests that the model is capable of learning and performing well on new, unseen data.

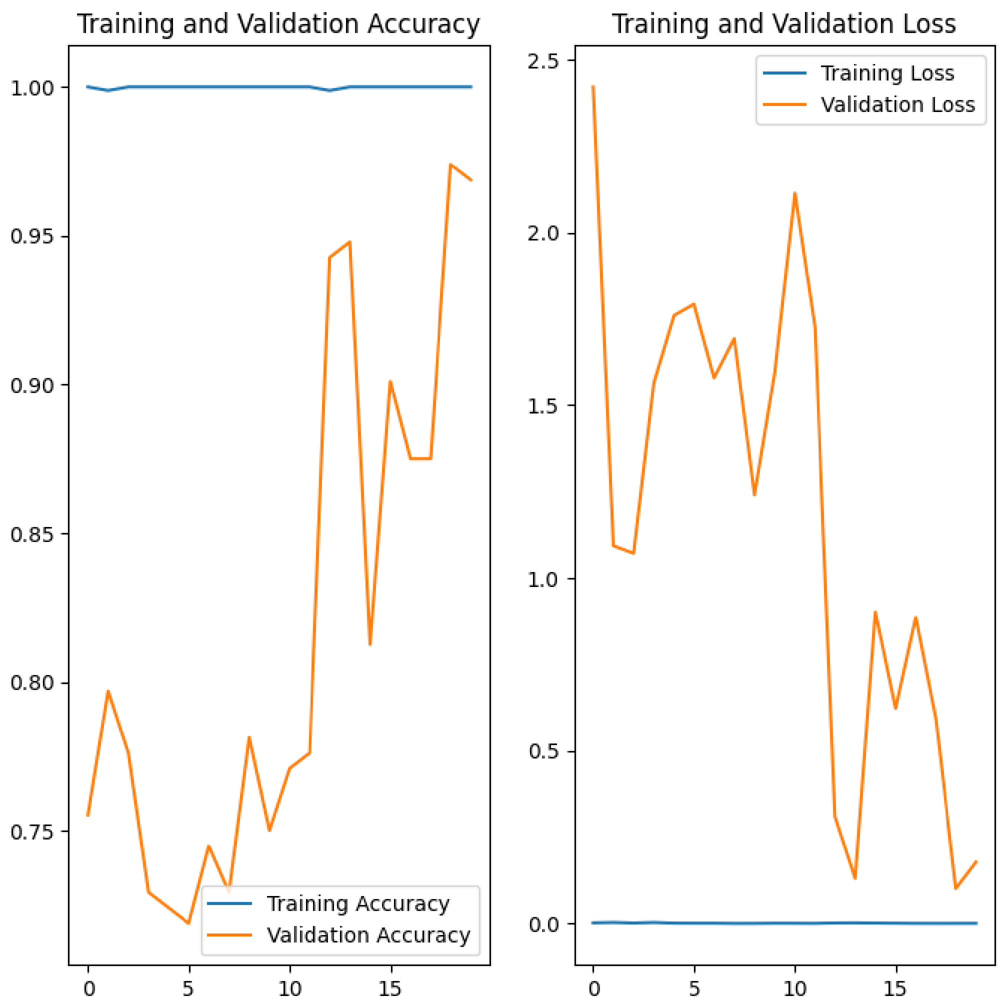

Figure 4.2.2: Accuracy and Loss Graph of CNN-02

Here is another graph of the CNN-02 model. The graph shows a decreasing trend for both the training and validation loss. This suggests that the model is learning from the training data and generalizing well to the validation data. Additionally, the high training accuracy and high validation accuracy shows that the model is learning and predicting the correct target value on training and validation data and improving over time.

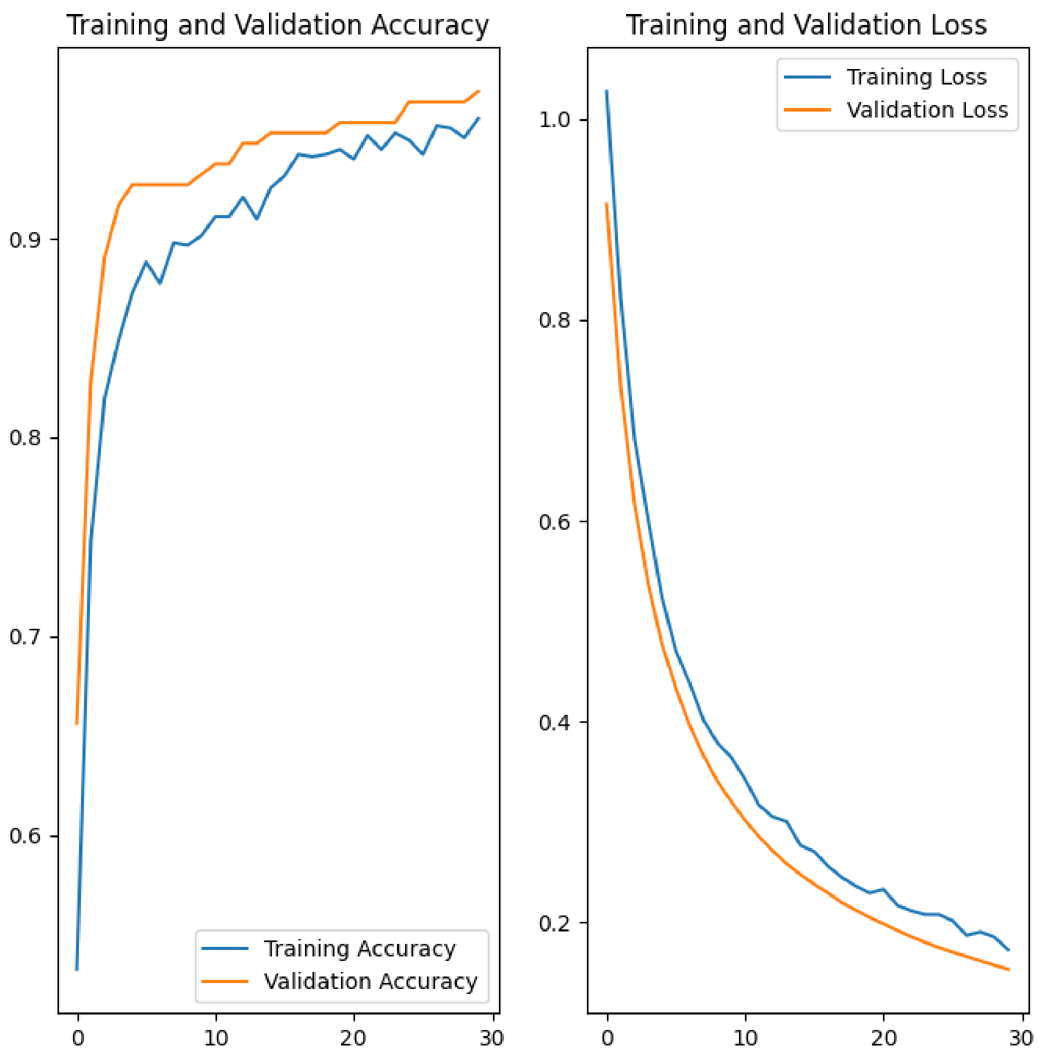

Figure 4.2.3: Accuracy and Loss Graph of EfficientNetB4

Analyzing the graph of EfficientNetB4 in figure 4.2.3, there are several factors to be noted. Initially, both training and validation accuracy increase at a relatively fast rate, indicating high initial learning capacity. The validation accuracy then stops increasing due to the model reaching a high level of accuracy and being unable to learn more from the validation dataset. Similarly, the training accuracy plateaus at around 90%, indicating that the model has also learned as much as it can from the training dataset. The training and validation loss, on the other hand, decrease over time, showing the model's learning progress. However, the final accuracy values of the model, which are close to 1.0, indicate that the model has achieved very high accuracy in image classification. Overall, this EfficientNet-B4 model has performed well in image classification, achieving a high level of accuracy with an efficient learning process.

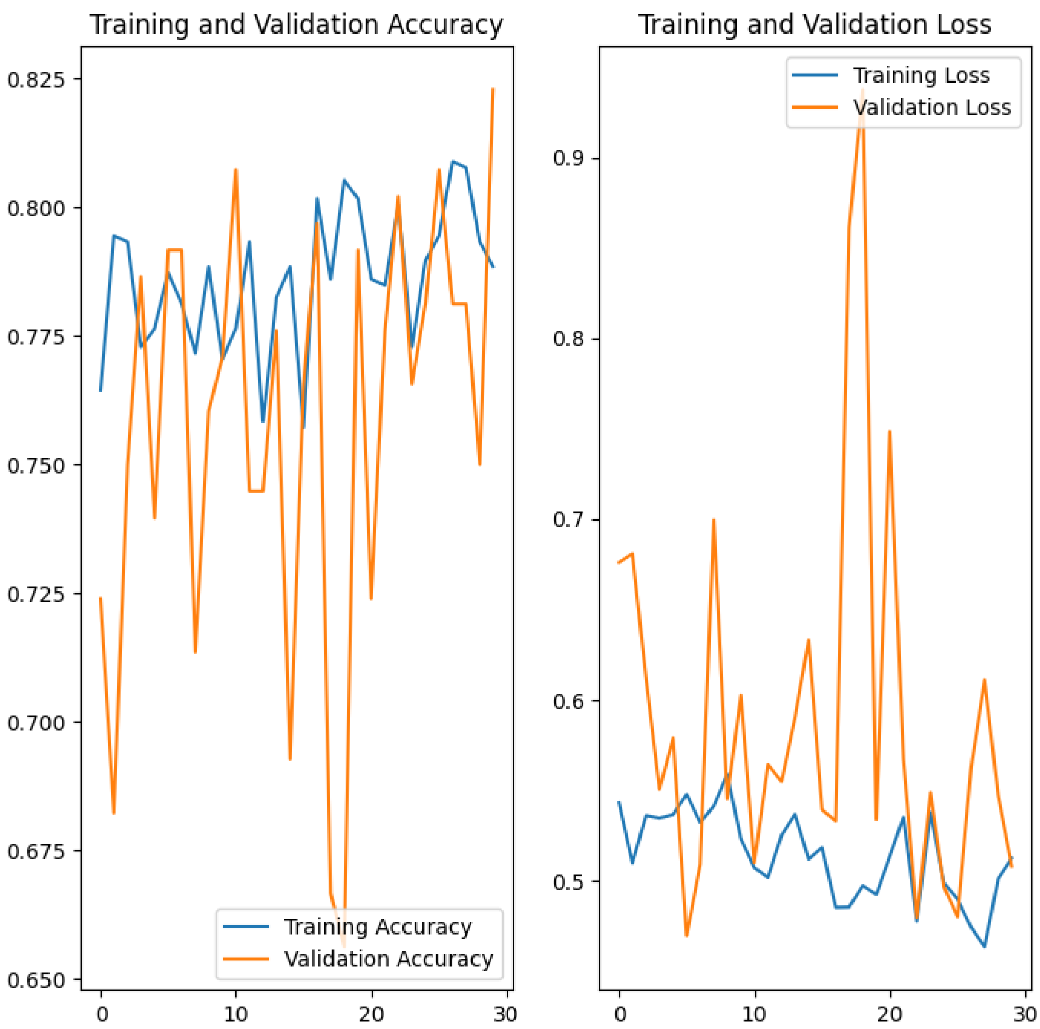

Figure 4.2.4: Accuracy and Loss Graph of InceptionV3

The graph of the InceptionV3 model in figure 4.2.4 shows the improvement of models training performance. Training accuracy increases from around 0.6 to 0.9, while validation accuracy rises from around 0.6 to 0.8. The decrease in both training and validation loss indicates that the model is learning and improving its predictions. Overall, the performance of the model can be considered satisfactory but not impressive, as both the training and validation accuracy increase significantly over time. The model is capable of learning complex patterns and improving its predictions.

Confusion Matrix

To evaluate a model from its confusion matrix is the best practice for an overall result exhibition. It is expected from a model to produce high accuracy in both positive and negative predictions, minimizing mistakes like false positives and false negatives. Precision, recall, and F1-score provide more detailed insights into specific aspects of the model's performance. Specific goals and requirements of certain task must be considered to interpret the result.

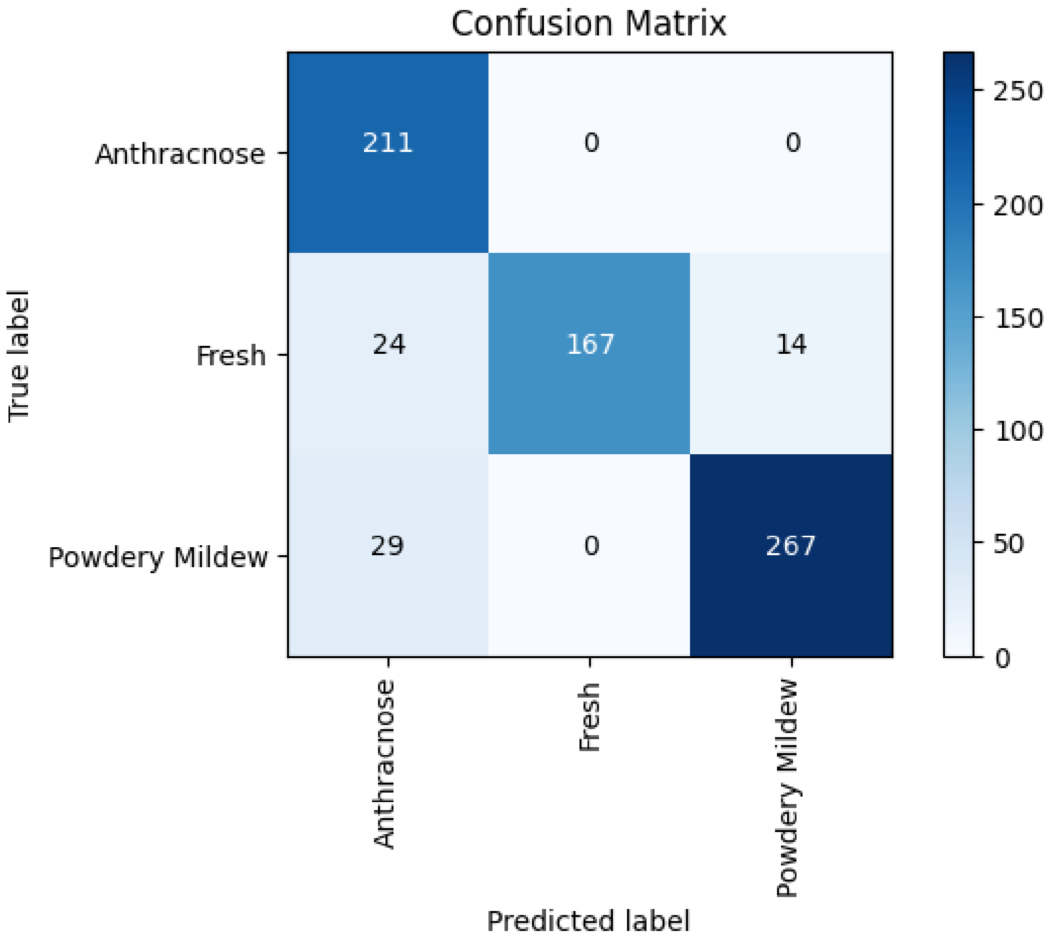

Figure 4.2.5: Confusion Matrix of CNN-01

If we analyze the classification result of the proposed CNN-01 model in figure 4.2.5, we can see the confusion matrix providing outstanding prediction result in properly classifying the diseases. The model produces an accuracy of 90.59% with the precision, recall and f1 score of 0.93, 0.90 & 0.9 respectively. The result indicates a commendable result in overall.

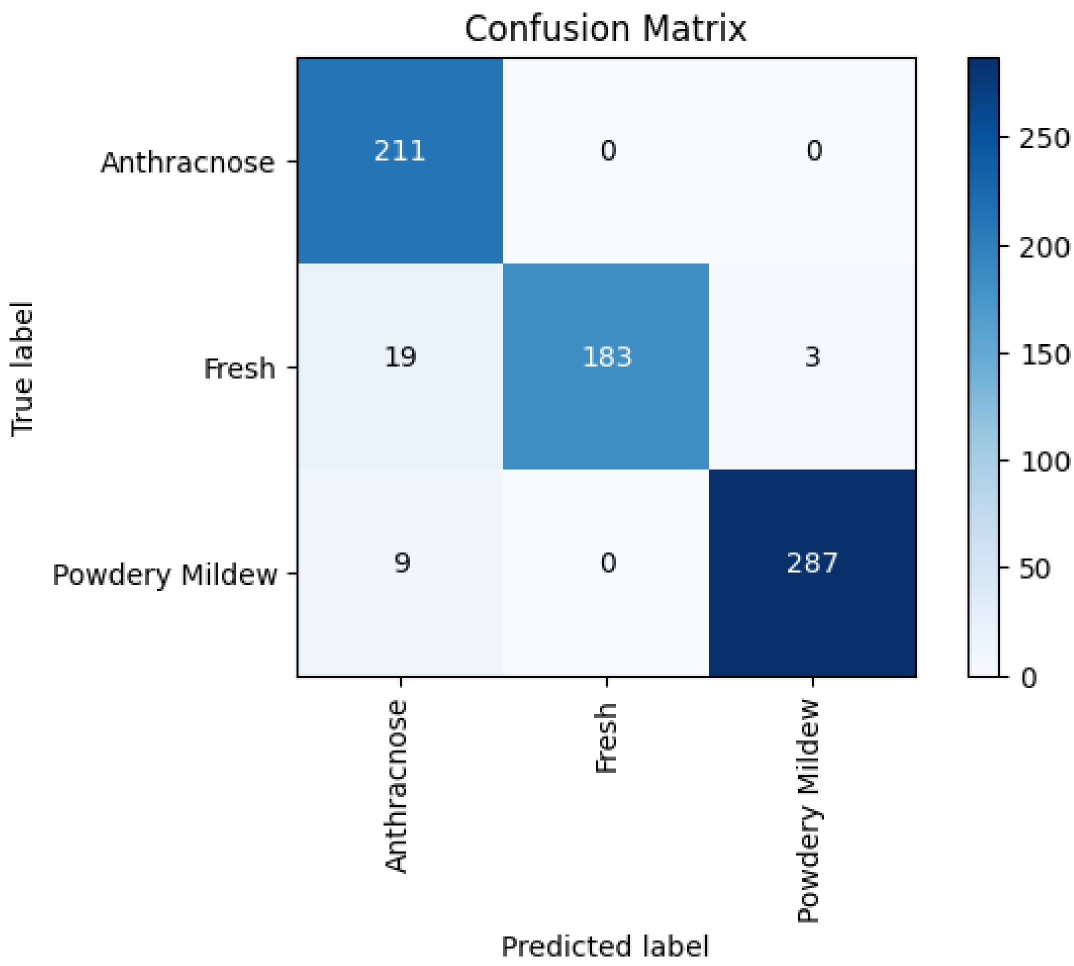

Figure 4.2.6: Confusion Matrix of CNN-02

Inspecting the classification report from the confusion matrix in figure 4.2.6, the proposed CNN-02 model generates highest accuracy among all the models, generating 95.65% accuracy, with the precision, recall and f1 score of 0.953, 0.967 & 0.958 respectively. The result indicates better performance than the CNN model without batch normalization and dropout layers.

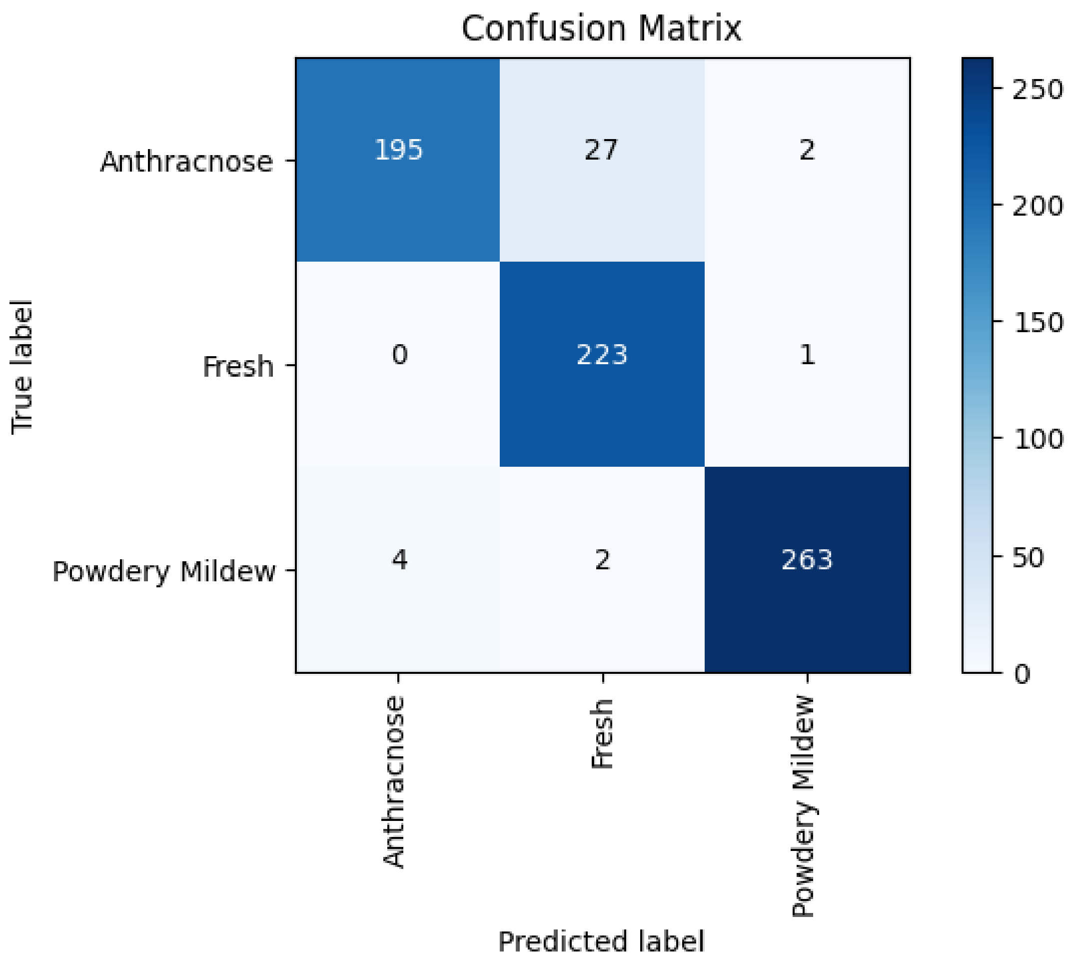

Figure 4.2.7: Confusion Matrix of EfficientNetB4

If we analyze the classification result of the proposed EfficientNetB4 model in figure 4.2.7, we can see the confusion matrix providing outstanding prediction result in properly classifying the diseases. The model generates precision, recall and f1-score of 0.983, 0.947 & 0.964 respectively. The overall accuracy of the model is 94.98%, which is quite remarkable for a pre-trained model. The result also shows an interesting fact that, the model generated highest precision rate among all the models in this experiment, thanks to the efficient learning process of EfficientNetB4.

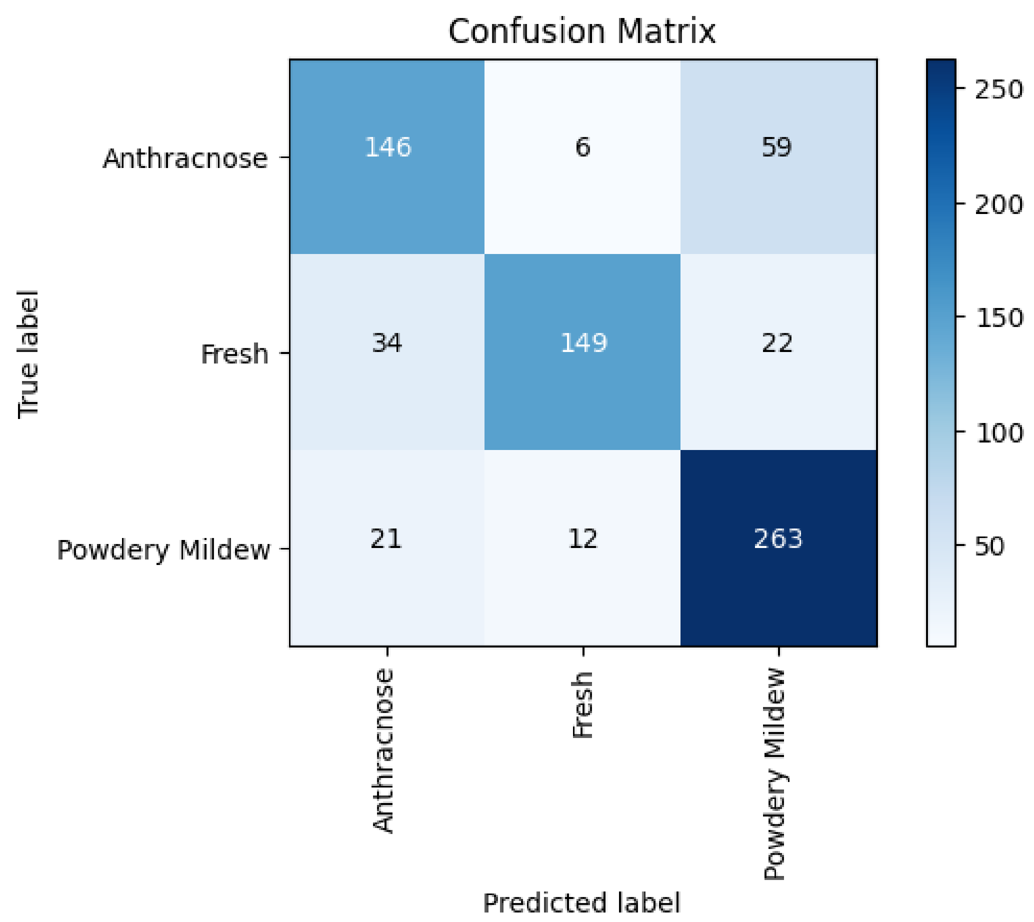

Figure 4.2.8: Confusion Matrix of InceptionV3

The confusion matrix of InceptionV3 model from figure 4.2.8 shows fluctuations in data predictions in incorrect class identification. Overall, the confusion matrix generates a poor accuracy of 76%. Compared to all other models, InceptionV3 model was the least recommendable model in terms of the given data prediction.

The overall comparison of the models from the table 4.2.1 implies that the highest result was achieved by CNN-02 model, thanks to its improved structure over CNN-01 that allowed the model to have an accuracy of 95.65%.

TABLE 4.2.1: CLASSIFICATION RESULTS

Model Name	Accuracy (%)	Precision	Recall	F1-Score
CNN-01	90.59%	0.93	0.90	0.90
CNN-02	95.65%	0.953	0.967	0.958
EfficientNetB4	94.98%	0.983	0.947	0.964
InceptionV3	76.01%	0.764	0.746	0.749

4.3 Discussion

Even though the research is a comparative analysis of deep learning models, one thing we must note is that, every model handles images in its own characteristics. Here the first CNN model initially learns slowly, but over time, it adapts to the unseen data and produces better prediction results. The modified CNN on the other hand, produces more accurate results than CNN-01 in case of learning and validating data, thanks to its batch normalization and dropout layers that produced the highest accuracy in this experiment.

Moving on to the transfer learning, pretrained EfficientNetB4 had high initial learning capacity with efficient learning process, which was able to predict the correct classes with the highest precision rate among the models. But in the case of InceptionV3 model performed poor on the dataset, even though it was capable of identifying complex pattern and improving its learning process, but the model couldn't reach its highest learning capability and produced a lower result compared to others. Adding more images to the dataset could've made the model worthy in the comparison. The overall accuracy result comparison is shown in figure 4.3.1.

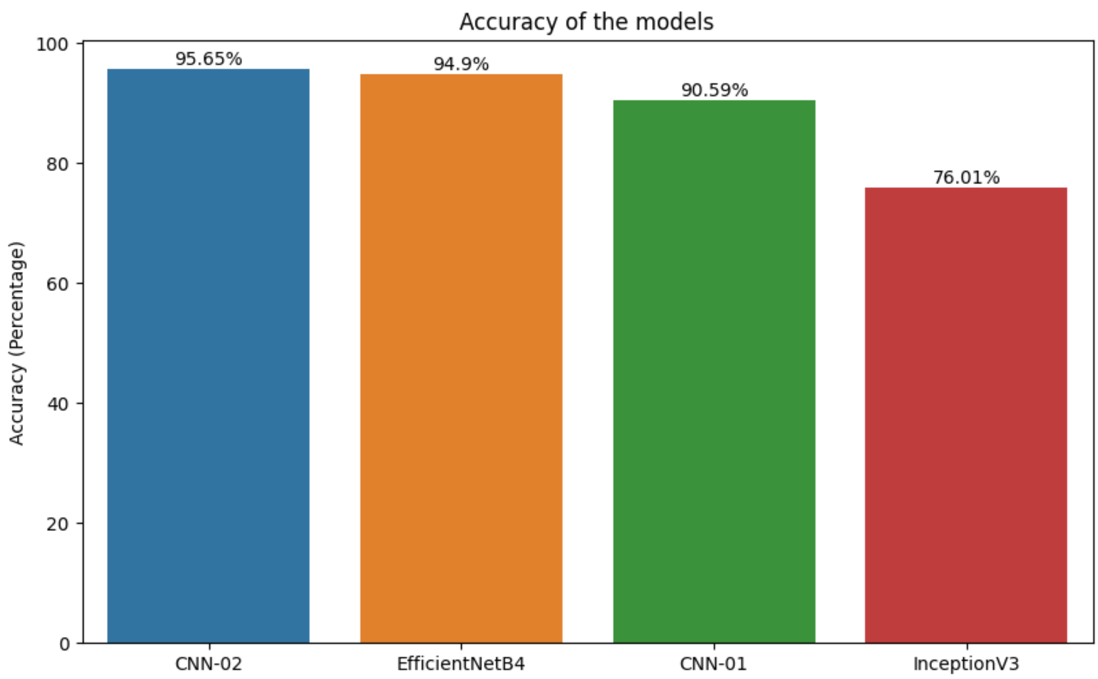

Figure 4.3.1.: Accuracy Result Comparison

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

The incorporation of Convolutional Neural Networks (CNN) and transfer learning in the detection of mango leaf diseases is a significant breakthrough in agricultural technology. This novel methodology, which I consider to be a paradigm shift from conventional methodologies, presents a more rigorous and evidence-based approach to disease detection. This technology enables the automation and improvement of disease identification, resulting in more precise and timely interventions. As a result, it can greatly reduce crop damage and loss of output which is particularly crucial in areas where mango cultivation is the mainstay of rural economies and serves as a significant agricultural endeavor. Furthermore, my analysis indicates that the utilization of this sophisticated technology in agriculture yields significant socioeconomic advantages since we all know that enhanced disease detection benefits small-scale and large-scale farmers by promoting crop health, boosting yields, and potentially enhancing financial stability and it eventually leads to enhancement of quality of life for entire agricultural communities. In addition, maintaining a stable mango supply through constant crop health can help control market prices and ensure food availability. Moreover, the increasing demand for skilled individuals in artificial intelligence and deep learning specifically in the field of agriculture indicates the possibility of work opportunities in the technology industry, which is a development that I find particularly encouraging.

5.2 Impact on Environment

After doing a lot of research, I've learned that one of the biggest environmental benefits of using CNN and transfer learning to find diseases on mango leaves is that pesticides and fungicides might not have to be used as much. By identifying sick plants, farmers can apply treatments more precisely, saving resources and reducing the amount of chemical runoff that enters nearby ecosystems. This targeted method is very important for protecting biodiversity, soil health, and water quality. Also, this technology is very important for keeping the environment in balance, which is in line with the ideas of sustainable agriculture, which I highly support. Making it easier to find and treat diseases early on, it helps keep crops healthy so they can handle changes in the weather with less need for chemical help. Over time, this makes mango growing much more sustainable, making sure that it will continue to be an environmentally friendly method for many years to come.

5.3 Ethical Aspects

Even though growth in agricultural technology is good, it brings up important moral issues. To protect farmers' information, I think it is very important to strictly follow data privacy rules when gathering and analyzing data through CNN and transfer learning. To make sure that everyone gets an equal share of the benefits, this technology needs to be easy for all farms, no matter how much money they have, to access and afford. For this to happen, small-scale farmers need to get the training and help they need because they could be left out or hurt by the digital change. Getting the right mix of technology and custom. Also, I think it's important to find a balance between accepting new tools and sticking to traditional farming methods. When Deep Learning is used in agriculture, it should work with the vast information and methods that have been passed down through generations of farmers, not replace them.

5.4 Sustainability Plan

I believe that ongoing study and development should be a part of a complete long-term plan for using CNN and transfer learning to find diseases in mango leaves. Not only does this mean making the technology better, but it also means adapting it to different illnesses and environmental conditions. Innovation that lasts is important for dealing with new diseases and adapting to changing weather conditions. Also, training programmers are very important for making sure that this technology is used correctly. Customized training programmers for farmers, farm workers, and local development services are very important for making the technology easier to understand and use. Additionally, forming partnerships with governmental bodies, agricultural institutions, and non-governmental organizations is necessary to include these technologies in broader farming policies and methods, which will ensure their responsible and effective use. Besides that, it is important to track the performance of these technologies on a regular basis to see how they affect food health, farmer income, environmental protection, and the long-term viability of agriculture as a whole. Which means, we need to not only see how well the technology works but also understand its effects on society, the economy, and the environment over time. Lastly, feedback can be very important for figure out how to make things better and decide on the right rules and decisions. This can ensure that the technology continues to benefit farms, consumers, and the environment.

CHAPTER 6

SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH

6.1 Summary of the Study

Inspired from the need of efficient and accurate method, the research dives into the domain of leaf disease detection using deep learning. The raw data containing diverse images of diseases are utilized to the maximum, thanks to the neural network-based models, the research successfully ensures comprehensive presentation of real-world situation. The experimental results demonstrate the success of the proposed methodologies in effectively identifying and classifying mango leaf diseases. The CNN models showcased their capability to learn unique patterns and features directly from the input data, achieving admirable accuracy rates. The comparative analysis highlights the performance of the models, differentiating the strengths and weaknesses. The study successfully concludes that both CNN and Transfer learning approaches can be the viable solution for leaf disease detection, contributing to the growing body of knowledge in agricultural technology. However, the choice between these methodologies may vary on application.

6.2 Conclusions

Starting from the idea of the research to the result acquisition, the study involves tremendous hard work and motivation. I have collected raw image data from local orchards of Rajshahi, which was augmented to improve the model training efficiency. According to the identification of leaf images, the data was divided into 3 classes which were normal mango leaves, anthracnose and powdery mildew. Then the data were used to train some models, including convolutional neural network, EfficientNetB4 and InceptionV3. Among them, CNN-02 outperformed other models in terms of learning improvement & better generalization result over time, but EfficientNetB4 with its efficient learning capability, didn't fall behind too far. The successful application of the models to previously unseen data underscores their potential for practical deployment in real-world agricultural scenarios contributing to the evolving field of agricultural technology. Even though more data could be fed to the models for the highest training efficiency, there still remains some drawbacks in the research. Not to mention, the research still holds tremendous possibilities for further success.

6.3 Implication for Further Study

Deep learning models for mango leaf disease prediction can revolutionize various sectors, particularly agriculture and horticulture. Their integration into agricultural farms and orchards enables early disease detection, helping farmers to implement timely and effective disease control measures. This technology is also invaluable in agricultural research and development, aiding in studying disease patterns and contributing to breeding more resilient mango varieties. The advent of precision agriculture, combining deep learning with drones or IoT, allows for meticulous monitoring of vast orchards, offering detailed plant health insights. Additionally, agricultural consultancy services can leverage these models to provide advanced disease management advice. Government agriculture departments might find these models helpful for regional or national crop health monitoring, influencing policy and resource allocation for disease control.

REFERENCES

[1] Agriculture Overview: Development News, Research, Data | World Bank, available at <<https://www.worldbank.org/en/topic/agriculture/overview>>, last accessed on 02-01-2024 at 14:29 PM.

[2] Impact of Sustainable Agriculture and Farming Practices, available at <<https://www.worldwildlife.org/industries/sustainable-agriculture>>, last accessed on 02-01-2024 at 02:30 PM.

[3] Chia Seeds, 7 Health Benefits, available at <<https://www.healthline.com/nutrition/11-proven-health-benefits-of-chia-seeds?fbclid=IwAR2rTNASqeV9D0h6j-p2Q7O9PnE-OrvThQwaPMGupt8B3iYtMmw04-qLuK0#TOC_TITLE_HDR_13>>, last accessed on 02-01-2024 at 02:32 PM.

[4] Mango Production by Country 2024, available at <<https://worldpopulationreview.com/country-rankings/mango-production-by-country>>, last accessed on 02-01-2024 at 02:33 PM.

[5] Suljović, Almira, et al. "Detection of Plant Diseases Using Leaf Images and Machine Learning." *2022 21st International Symposium INFOTEH-JAHORINA (INFOTEH)*. IEEE, 2022.

[6] Mohanty, Sharada P., David P. Hughes, and Marcel Salathé. "Using deep learning for image-based plant disease detection." *Frontiers in plant science* 7 (2016): 1419.

[7] Liu, Bin, et al. "Identification of apple leaf diseases based on deep convolutional neural networks." *Symmetry* 10.1 (2017): 11.

[8] Lu, Yang, et al. "Identification of rice diseases using deep convolutional neural networks." *Neurocomputing* 267 (2017): 378-384.

[9] Singh, Uday Pratap, et al. "Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease." *IEEE access* 7 (2019): 43721-43729.

[10] Arivazhagan, S., and S. Vineth Ligi. "Mango leaf diseases identification using convolutional neural network." *International Journal of Pure and Applied Mathematics* 120.6 (2018): 11067-11079.

[11] Merchant, Mustafa, et al. "Mango leaf deficiency detection using digital image processing and machine learning." *2018 3rd International Conference for Convergence in Technology (I2CT)*. IEEE, 2018.

[12] Chen, Jing, Qi Liu, and Lingwang Gao. "Visual tea leaf disease recognition using a convolutional neural network model." *Symmetry* 11.3 (2019): 343.

[13] Agarwal, Mohit, et al. "ToLeD: Tomato leaf disease detection using convolution neural network." *Procedia Computer Science* 167 (2020): 293-301.

[14] Aravind, Krishnaswamy R., et al. "Disease classification in Solanum melongena using deep learning." *Spanish Journal of Agricultural Research* 17.3 (2019): e0204-e0204.

[15] Arya, Sunayana, and Rajeev Singh. "A Comparative Study of CNN and AlexNet for Detection of Disease in Potato and Mango leaf." *2019 International conference on issues and challenges in intelligent computing techniques (ICICT)*. Vol. 1. IEEE, 2019.

[16] Trang, Kien, et al. "Mango diseases identification by a deep residual network with contrast enhancement and transfer learning." *2019 IEEE conference on sustainable utilization and development in engineering and technologies (CSUDET)*. IEEE, 2019.

[17] Nagaraju, Y., et al. "Transfer learning based convolutional neural network model for classification of mango leaves infected by anthracnose." *2020 IEEE International Conference for Innovation in Technology (INOCON)*. IEEE, 2020.

[18] Maheshwari, Kanak, and Amit Shrivastava. "A review on mango leaf diseases identification using convolution neural network." *International Journal of Scientific Research & Engineering Trends* 6.3 (2020): 1399-1403.

[19] Kumar, Pankaj, et al. "Classification of mango leaves infected by fungal disease anthracnose using deep learning." *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, 2021.

[20] Rajbongshi, Aditya, et al. "Recognition of mango leaf disease using convolutional neural network models: a transfer learning approach." *Indonesian Journal of Electrical Engineering and Computer Science* 23.3 (2021): 1681-1688.

[21] Sharma, Amisha, et al. "Mango Leaf Diseases Detection using Deep Learning." *International Journal of Knowledge Based Computer Systems* 10.1 (2022).

[22] Jyothi, E. V. N., and M. Kranthi. "Application of Transfer Learning to Convolutional Neural Network Models for Mango Leaf Disease Recognition."

[23] Rizvee, Redwan Ahmed, et al. "LeafNet: A proficient convolutional neural network for detecting seven prominent mango leaf diseases." *Journal of Agriculture and Food Research* 14 (2023): 100787.

[24] Mimi, Afsana, et al. "Identifying selected diseases of leaves using deep learning and transfer learning models." *Machine Graphics & Vision* 32.1 (2023).

[25] Vijay, C. P., and K. Pushpalatha. "Revolutionizing Mango Leaf Disease Detection: Leveraging Segmentation and Hybrid Deep Learning for Enhanced Accuracy and Sustainability." *International Journal of Intelligent Systems and Applications in Engineering* 11.4 (2023): 121-131.

PLAGIARISM REPORT

A COMPARATIVE ANALYSIS OF CNN AND TRANSFER LEARNING APPROACHES FOR MANGO LEAF DISEASE DETECTION

ORIGINALITY REPORT

15%	11%	8%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	3%
2	Submitted to Daffodil International University Student Paper	1%
3	Submitted to CSU Northridge Student Paper	1%
4	www.coursehero.com Internet Source	1%
5	Submitted to College of Banking and Financial Studies Student Paper	1%
6	dokumen.pub Internet Source	<1%
7	"Third International Conference on Image Processing and Capsule Networks", Springer Science and Business Media LLC, 2022 Publication	<1%
8	Submitted to Landmark University Student Paper	<1%